

KLASIFIKASI MINAT SISWA UNTUK PROGRAM STUDI JURUSAN TEKNOLOGI INFORMASI - POLITEKNIK NEGERI SAMARINDA MENGGUNAKAN METODE FUZZY C-MEANS CLUSTERING

Bedi Suprpty¹, Fariyanti²

¹Jurusan Teknologi Informasi, Politeknik Negeri Samarinda,
Jl. Cipto Mangunkusumo Kampus Gunung Lipan Samarinda - Kalimantan Timur
Email: Bedirheody@gmail.com

² Jurusan Akuntansi, Politeknik Negeri Samarinda
Jl. Cipto Mangunkusumo Kampus Gunung Lipan Samarinda - Kalimantan Timur
Email: fariyantismd@gmail.com

ABSTRAK

Siswa yang telah lulus sekolah menengah atas atau sekolah menengah kejuruan akan melanjutkan ke jenjang yang lebih tinggi seperti Politeknik Negeri Samarinda. Politeknik Negeri Samarinda terdiri dari beberapa Jurusan di antaranya adalah jurusan Teknologi Informasi. Jurusan Teknologi Informasi memiliki 4 Program Studi antara lain; D3 Teknik Informatika, D3 Teknik Komputer, D4 Teknologi Informasi Multimedia dan D4 Rekayasa Teknologi Komputer. Penelitian ini dilakukan dengan mengklasifikasikan peminatan siswa-siswi yang akan melanjutkan studi ke Jurusan Teknologi Informasi Politeknik Negeri Samarinda. Data sumber diperoleh dari kuesioner. Pengumpulan data dilakukan dengan metode kuesioner, kuesioner terdiri dari 15 butir pertanyaan dan mempunyai 5 kriteria. Masing-masing kriteria mempunyai 3 butir pertanyaan. Kuesioner disebar kepada 160 siswa-siswi Sekolah Menengah Atas dan Sekolah Menengah Kejuruan yang ada di kota Samarinda Kalimantan Timur. Cluster pada penelitian ini terbagi menjadi 4, yaitu cluster 1 minat terhadap Program Studi D3 Teknik Informatika, cluster 2 minat terhadap Program Studi D3 Teknik Komputer, cluster 3 minat terhadap Program Studi D4 Teknologi Informasi Multimedia dan cluster 4 minat terhadap D4 Teknologi Rekayasa Komputer. Metode Fuzzy C-Means digunakan dalam menyelesaikan permasalahan tersebut dimana hasil pengelompokan cluster 1 beranggotakan 41 siswa-siswi, cluster 2 beranggotakan 46 siswa-siswi, cluster 3 beranggotakan 21 siswa-siswi, cluster 4 beranggotakan 52 siswa-siswi. Rata-rata persentase MAPE untuk keseluruhan cluster sebesar 27,07%.

Kata kunci: Sekolah Menengah Umum (SMU), Sekolah Menengah Kejuruan (SMK), Clustering, Fuzzy C-means, MAPE.

ABSTRACT

Students who have graduated from high school or high school will continue to a higher level such as Samarinda State Polytechnic. Samarinda State Polytechnic consists of several Departments approved by the Information Technology department. The Information Technology Department has 4 Study Programs including; D3 Informatics Engineering, D3 Computer Engineering, D4 Multimedia Information Technology and D4 Computer Technology Engineering. This research was conducted by classifying specialization of students who would continue their studies to the Department of Information Technology, Samarinda State Polytechnic. Sources of data obtained from the questionnaire. Data collection was carried out by questionnaire method, the questionnaire consisted of 15 questions and had 5 criteria. Each criterion has 3 questions. The questionnaire was distributed to 160 high school and vocational high school students in the city of Samarinda, East Kalimantan. Clusters in this study are divided into 4, namely cluster 1, an interest in the D3 Study Program in Informatics Engineering, cluster 2 an interest in the D3 Study Program in Computer Engineering, cluster 3 an interest in the D4 Study Program in Multimedia Information Technology and cluster 4 an interest in D4 in Computer Engineering Technology. Fuzzy C-means method is used in resolving these complications where the results of clustering cluster 1 consists of 41 students, cluster 2 consists of 46 students, cluster 3 consists of 21 students, cluster 4 consists of 52 students. The average MAPE percentage for the whole cluster is 27.07%.

Keywords: General High School (SMU), Vocational High School (SMK), Clustering, Fuzzy C-means, MAPE.

1. PENDAHULUAN

Siswa-siswi yang telah lulus dari Sekolah Menengah Umum (SMU) atau Sekolah Menengah Kejuruan (SMK) akan melanjutkan jenjang yang lebih tinggi antara lain; Universitas, Sekolah Tinggi dan Politeknik. Politeknik Negeri Samarinda adalah salah satu jenjang yang bisa ditempuh setelah lulus SMU atau SMK. Politeknik Negeri Samarinda khususnya Jurusan Teknologi Informasi dibagi atas 4 Program Studi antara lain; D3 Program Studi Teknik Informatika, D3 Teknik Komputer, D4 Teknologi Informasi Multimedia dan D4 Teknologi Rekayasa Komputer.

Jurusan Teknologi informasi Sekolah menengah atas (SMA) memiliki tujuan untuk mempersiapkan siswa memasuki perguruan tinggi yang berkaitan dengan ilmu-ilmu program vokasi untuk siap memasuki lapangan kerja, menyiapkan tenaga kerja tingkat menengah dan mengisi kebutuhan dunia usaha, menyiapkan mahasiswa lulusan agar mampu berkarir. Untuk itu penelitian ini dilakukan untuk mengetahui minat siswa siswi untuk sekolah menengah umum dan sekolah menengah kejuruan di kelompokkan ke dalam beberapa *cluster* yang sesuai dengan minat ke ilmu yang terkait dengan program studi pada Jurusan Teknologi Informasi Politeknik Negeri Samarinda.

Peneliti sebelumnya yang dilakukan oleh Karti (2013) melakukan penelitian dengan mengelompokkan kabupaten dan kota di Jawa Timur berdasarkan tingkat kemiripan berdasarkan indikator partisipasi pendidikan SMA/SMK/MA menggunakan metode fuzzy c-means [1].

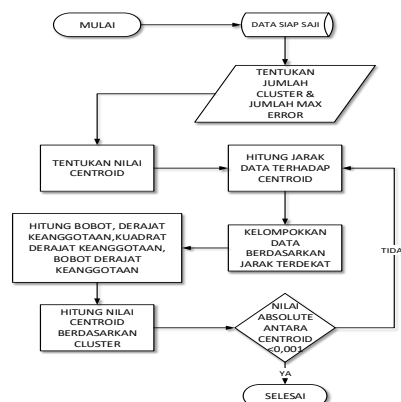
Clustering adalah proses untuk mengklasifikasi obyek ke dalam kelompok-kelompok yang relative didasarkan obyek di dalam setiap kelompok harus relative mirip [2]. Ada berbagai macam metode *clustering* antara lain adalah K-means dan Fuzzy C-means. Dalam penelitian ini metode yang dipilih adalah metode Fuzzy C-means. Fuzzy C-means (FCM) adalah salah satu teknik pengelompokan data dimana masing-masing keberadaan tiap titik data dalam suatu *cluster* ditentukan oleh derajat keanggotaan [3][9].

Penelitian ini dilakukan untuk mengetahui klasifikasi Program Studi berdasarkan minat dari siswa-siswi SMU dan SMK. Dimana pembagian *cluster* pada penelitian ini dibagi menjadi 4 bagian, *cluster 1* berisi siswa siswi SMU dan SMK yang mempunyai minat terhadap D3 Program Studi Teknik Informatika, *cluster 2* berisi siswa siswi SMU dan SMK yang mempunyai minat terhadap D3 Teknik Komputer, *cluster 3* berisi siswa siswi SMU dan SMK yang mempunyai minat terhadap D4 Teknologi Informasi Multimedia dan yang terakhir adalah *cluster 4* berisi siswa siswi sekolah menengah pertama yang mempunyai minat terhadap D4 Teknologi Rekayasa Komputer. Pada hasil akhir, akan diperoleh klasifikasi sekolah menengah atas yang terbagi menjadi beberapa *cluster*.

2. MATERI DAN METODE

A. Analisis Cluster

Analisis cluster adalah proses pengelompokan data ke dalam kelas (*cluster*), dengan tujuan agar objek dalam satu *cluster* memiliki kemiripan yang lebih tinggi dibandingkan objek *cluster* lain. Ketidak miripan didasarkan pada nilai atribut yang menggambarkan suatu objek. Objek-objek dengan jarak yang lebih kecil diantara mereka adalah lebih serupa dibandingkan objek-objek dengan jarak yang lebih besar [4]. ED (*Euclidean Distance*) adalah salah satu teknik pengukuran jarak. Teknik pengukuran ini dipergunakan dalam clustering karena sederhana tetapi sangat *sensitive*. Formula perhitungan ED ditunjukkan pada persamaan 1 [5].



Gambar 1. Diagram alir Fuzzy C-means.

$$d = \sqrt{(\chi_r - \chi_s)^2 + (\chi_r - \chi_s)^2} \quad (1)$$

B. Fuzzy C-Means

Ada beberapa algoritma *clustering* data, salah satunya adalah Fuzzy C-means (FCM). FCM adalah suatu teknik *cluster* yang keberadaan tiap-tiap titik data suatu *cluster* ditentukan oleh derajat keanggotaan. Teknik ini pertama kali

dikenalkan oleh Jim Bezdek pada tahun 1981 [6][7]. Kelebihan dari metode Fuzzy C-means dapat melakukan *clustering* lebih dari *variable* secara sekaligus [8]. Konsep dasar FCM seperti yang ditunjukkan pada gambar 1.

1) *Data Siap Saji*

Data sudah siap yang dimaksud adalah data yang telah dianalisa dan dilakukan normalisasi untuk di implementasikan metode fuzzy c-means.

2) *Menentukan jumlah cluster & jumlah max error*

Pada tahapan ini akan dilakukan penentuan jumlah *Cluster* yang akan dibangun sesuai dengan studi kasus dalam penelitian ini yaitu berjumlah 4 *cluster*. Data yang lain yang dibutuhkan dalaha menentukan jumlah *maximun error* yaitu sebesar 0,0001.

3) *Menentukan nilai centroid*

Tahapan ini dilakukan dengan cara membuat suatu formulasi untuk menentukan centroid awal menggunakan teknik mencari *interval* data *minimum* dan *maximun* yang kemudian dibagi berdasarkan jumlah *Cluster* secara sistematis. adapun tahapan sebagai berikut ;

a) Menentukan bobot masing-masing atribut dengan persamaan sebagai berikut :

$$\text{Bobot Nilai} = \text{Nilai Max} - \text{Nilai Min} \quad (2)$$

b) Menghitung interval kelas dengan persamaan sebagai berikut :

$$\text{Interval kelas} = \frac{\text{Bobot Nilai}}{\text{Banyaknya Kelas}} \quad (3)$$

c) Mengambil nilai tengah dari masing-masing setiap kelas disetiap atribut untuk dijadikan nilai centroid awal.

4) *Menghitung jarak data terhadap centroid*

Untuk mendapatkan nilai ED (*Euclidean Distance*) dengan cara masing-masing atribut dikurang centroid awal lalu dikuadratkan. Tambahkan semua, hasil setelah itu diakarkan. Rumus perhitungan ED ditunjukkan pada persamaan 1.

5) *Mengelompokkan data berdasarkan jarak terdekat*

Untuk mendapatkan hasil pengelompokkan data, dengan cara mencari nilai terkecil dari setiap *cluster* yang mendekati nilai centroid.

6) *Proses clustering*

Proses *clustering* metode fuzzy c-means keberadaan data dalam suatu kelas atau *cluster* yang ditentukan oleh derajat keanggotaan.

a) Pembobot ED

Untuk menghitung pembobotan hasil perhitungan jarak menggunakan *Euclidean distance* (ED). Dengan cara 1 dibagi dengan hasil perhitungan jarak. formula perhitungan bobot ED ditunjukkan pada persamaan 4.

$$\text{Bobot ED} = 1/\text{ED} \quad (4)$$

Keterangan :

ED = *Euclidean distance*

b) Derajat Keanggotaan

Derajat keanggotaan adalah bobot fuzzy. Nilai derajat keanggotaan didapat dengan cara membagi bobot *Euclidean distance* dengan total bobot masing-masing data. Formula perhitungan derajat keanggotaan ditunjukkan pada persamaan 5.

$$U = \frac{\text{Bobot ED}}{\text{Total bobot ED}} \quad (5)$$

Keterangan :

U = derajat keanggotaan

c) Kuadrat Derajat Keanggotaan

Dengan cara nilai setiap derajat keanggotaan dikuadratkan. Formula perhitungan kuadrat derajat keanggotaan ditunjukkan pada persamaan 6.

$$U = U^2 \quad (6)$$

Keterangan :

U = derajat keanggotaan

d) Bobot Derajat Keanggotaan

Untuk mendapatkan pembentukan centroid baru dihitung dengan menggunakan bobot derajat keanggotaan. Vektor 1 pada bobot derajat keanggotaan merupakan centroid pertama. Menghitung bobot derajat keanggotaan dengan cara Vektor 1 (V1) kuadrat derajat keanggotaan dikali dengan *feature vector* dan begitu juga dengan V2 sampai V4. Formula perhitungan bobot derajat keanggotaan ditunjukkan pada persamaan 7.

$$\text{Bobot derajat keanggotaan} = U^2 \times X \quad (7)$$

Keterangan :

X = feature vector

7) Pembangkitan Centroid Baru

Setelah mendapatkan hasil dari jumlah bobot derajat keanggotaan maka dapat ditentukan centroid baru yaitu dengan cara membagi jumlah masing-masing vektor pada kuadrat derajat keanggotaan dengan total kuadrat derajat keanggotaan untuk masing-masing vektor. Formula perhitungan pembangkitan centroid baru ditunjukkan pada persamaan 8.

$$\text{centroid baru} = \frac{\text{total } p}{\text{total } U^2} \quad (8)$$

Keterangan :

p = bobot derajat keanggotaan

U^2 = kuadrat derajat keanggotaan

8) Nilai Absolute antara Centroid < 0,00001

Untuk menghitung nilai max error dengan cara nilai centroid lama dikurangkan dengan centroid baru. Setelah itu nilai di totalkan jika nilai max error kurang dari 0,00001 maka iterasi selesai, dan sebaliknya jika nilai max error lebih dari 0,00001 maka iterasi dilanjutkan. Penghentian iterasi dilakukan ketika nilai *absolute* antara centroid sebelumnya dengan centroid baru lebih kecil dari 0,00001.

C. MEAN ABSOLUTE PERCENTAGE ERROR (MAPE)

MAPE merupakan rata-rata dari keseluruhan persentase kesalahan antara data actual dengan data hasil peramalan. MAPE dihitung dengan menggunakan kesalahan absolut pada tiap periode dibagi dengan nilai observasi yang nyata untuk periode itu. Kemudian merata-rata kesalahan persentase absolut tersebut. Pendekatan ini berguna ketika ukuran atau besar variable ramalan itu penting dalam mengevaluasi ketepatan ramalan. MAPE mengindikasikan seberapa besar kesalahan dalam meramal yang dibandingkan dengan nilai nyata [5].

$$MAPE = \frac{\sum \frac{X_t - F_t}{X_t} \times 100}{n} \quad (9)$$

Keterangan :

X_t = data aktual

F_t = hasil peramalan

N = banyak data

3. HASIL DAN PEMBAHASAN

A. Data Sumber

1) Pemilihan Responden

Responden yang dipilih dalam penelitian ini adalah siswa siswi Sekolah Menengah Umum (SMU) dan Sekolah Menengah Kejuruan (SMK) kota Samarinda - Kalimantan Timur. Responden berjumlah 160 orang di ambil secara acak pada setiap sekolah, yang berasal dari 2 kelas untuk setiap SMU dan SMK kota Samarinda.

2) Pemilihan Atribut

Pemilihan atribut untuk kuesioner di pilih berdasarkan pandangan secara umum siswa siswi untuk memilih sekolah yaitu : Atribut 1 : Akreditasi, Atribut 2 : Zona Sekolah, Atribut 3 : Biaya , Atribut 4 : Nem , Atribut 5 : Jarak Tempuh

3) Pembuatan Kuesioner

Kuesioner dibuat dalam format pertanyaan sebanyak 15 (sepuluh) butir. Untuk setiap atribut masing-masing memiliki 3 (tiga) pertanyaan. Bobot penilaian setiap pertanyaan ditunjukkan pada tabel 1.

Tabel 1. Bobot Penilaian

Bobot Penilaian	Skor
Tidak Penting	1
Kurang Penting	2
Cukup Penting	3
Penting	4
Sangat Penting	5

B. Analisa Data

Analisis data kuesioner adalah langkah untuk pemberian nilai pada kuesioner. Pemberian nilai dimaksud adalah untuk memperoleh total nilai dari setiap kelompok pertanyaan yang bertujuan untuk memperoleh nilai rata-rata dari setiap kelompok pertanyaan. Setelah kuesioner yang telah memperoleh nilai rata-rata kemudian di rekap berdasarkan

masing-masing kriteria. Setelah itu data sudah siap untuk di implementasi metode. Data responden ditunjukkan pada di tabel 2.

Tabel 2. Tabel Data Responden

No Responden	Atribut 1	Atribut 2	Atribut 3	Atribut 4	Atribut 5
1	4,000	3,333	2,667	2,000	3,667
2	3,667	3,667	4,333	3,000	3,000
3	2,000	2,333	3,000	3,000	3,667
4	3,333	3,333	2,333	3,000	2,000
5	3,000	4,000	2,333	4,667	3,333
..
156	4,667	3,000	4,000	3,333	4,333
157	3,000	2,333	2,333	3,000	2,333
158	4,000	3,667	3,333	1,667	3,333
159	2,667	3,667	3,333	3,000	2,000
160	2,000	2,333	2,667	3,333	2,333

C. Implementasi Metode Fuzzy C-means

1) Penentuan Nilai Centroid Awal

Setelah mendapatkan data di analisa dan siap yang ditunjukkan pada table 2, langkah selanjutnya adalah menentukan nilai centroid awal. Banyak jumlah *cluster* dalam penelitian ini adalah 4 *cluster* yang terdiri dari sebagai berikut : Program Studi D3 Teknik Informatika, Program Studi D3 Teknik Komputer, Program Studi D4 Teknologi Informasi Multimedia, Program Studi D4 Rekayasa Teknologi Komputer

Tahapan pertama adalah menentukan bobot masing-masing atribut dengan perhitungan berdasarkan persamaan 2.

Contoh perhitungan :

Diketahui :

Nilai $max = 5,000 \rightarrow$ untuk atribut 1

Nilai $min = 1,000 \rightarrow$ untuk atribut 1

Maka : Bobot nilai $= 5,000 - 1,000 = 3,000 \rightarrow$ untuk atribut 1

Setelah mendapatkan bobot nilai tahapan selanjutnya adalah menghitung interval kelas dengan perhitungan berdasarkan persamaan 3.

Contoh perhitungan :

Diketahui :

Bobot nilai $= 4,000 \rightarrow$ atribut 1

Banyak kelas $= 4$

Maka : Panjang interval kelas $= 4,000 / 4 = 1,000$

Berikut adalah hasil perhitungan untuk setiap atribut yang ditunjukkan pada tabel 3.

Tabel 3. Interval

Atribut 1	Bobot (Nilai Tertinggi - Nilai Terendah)	4,000
	Banyak Kelas	4,000
	Interval Kelas	1,000
Atribut 2	Bobot (Nilai Tertinggi - Nilai Terendah)	3,667
	Banyak Kelas	4,000
	Interval Kelas	0,917
Atribut 3	Bobot (Nilai Tertinggi - Nilai Terendah)	3,667
	Banyak Kelas	4,000
	Interval Kelas	0,917
Atribut 4	Bobot (Nilai Tertinggi - Nilai Terendah)	4,000
	Banyak Kelas	4,000
	Interval Kelas	1,000
Atribut 5	Bobot (Nilai Tertinggi - Nilai Terendah)	3,667
	Banyak Kelas	4,000
	Interval Kelas	0,917

Setelah nilai interval kelas didapat, langkah selanjutnya adalah mengambil nilai tengah dari masing-masing setiap kelas disetiap atribut untuk dijadikan nilai centroid awal. Berikut adalah perhitungan pengambilan nilai tengah pada atribut 1 :

Diketahui :

Nilai Max $= 5,000$

Nilai Min = 1,000
Interval kelas = 1,000
Maka :
Kelas 1 = 5,000 s/d 4,000
Kelas 2 = 4,000 s/d 3,000
Kelas 3 = 3,000 s/d 2,000
Kelas 4 = 2,000 s/d 1,000

Hasil Centroid Awal :

Nilai tengah pada kelas 1 adalah 4,500 dan nilai tersebut akan menjadi nilai centroid untuk *cluster* Sma Negeri pada atribut 1. Nilai centroid keseluruhan ditunjukkan pada tabel 4.

Tabel 4. Nilai Centroid Awal

<i>Cluster</i>	Atribut 1	Atribut 2	Atribut 3	Atribut 4	Atribut 5
D3 Teknik Informatika	4,500	4,208	4,542	4,167	4,208
D3 Teknik Komputer	3,500	3,292	3,625	3,167	3,292
D4 Teknologi Informasi Multimedia	2,500	2,375	2,708	2,167	2,375
D4 Teknik Rekayasa Komputer	1,500	1,458	1,792	1,167	1,458

D. Menghitung Jarak Data Terhadap Centroid

Setelah mendapatkan nilai centroid awal langkah selanjutnya adalah menghitung jarak data terhadap centroid yang berdasarkan persamaan 1. Hasil jarak data terhadap centroid ditunjukkan pada tabel 5.

Contoh Perhitungan :

Diketahui :

Data responden no. 1 : C1 = 4,000 → atribut 1
C2 = 3,333 → atribut 2
C3 = 2,667 → atribut 3
C4 = 2,000 → atribut 4
C5 = 3,667 → atribut 5

Nilai centroid awal K1 : C1 = 4,500 → atribut 1
C2 = 4,208 → atribut 2
C3 = 4,542 → atribut 3
C4 = 4,167 → atribut 4
C5 = 4,208 → atribut 5

Maka :

$$\sqrt{(4,000 - 4,500)^2 + (3,333 - 4,208)^2 + (2,667 - 4,542)^2 + (2,000 - 4,167)^2 + (3,667 - 4,208)^2} = 3.085$$

Tabel 5. Jarak Data Terhadap Centroid

No	Cluster 1	Cluster 2	Cluster 3	Cluster 4
V1	3,085	1,635	2,206	4,013
V2	1,963	0,885	2,599	4,666
V3	3,714	1,931	1,643	3,273
V4	3,639	1,842	1,609	3,290
V5	2,861	2,161	3,194	4,977
...
V156	1,578	1,643	3,457	5,499
V157	3,939	1,945	1,043	2,727
V158	3,003	1,651	2,340	4,156
V159	3,369	1,618	1,709	3,503
V160	4,182	2,243	1,271	2,691

E. Pengelompokan Data Berdasarkan Jarak Terdekat

Setelah mendapatkan nilai ED (*Euclidean Distance*), tahap selanjutnya adalah melakukan pengelompokan data berdasarkan jarak terdekat. Tabel pengelompokan data berdasarkan jarak terdekat dapat ditunjukkan pada tabel 6.

Tabel 6. Pengelompokan Data Berdasarkan Jarak Terdekat

No	Cluster 1	Cluster 2	Cluster 3	Cluster 4
V1	0	1	0	0
V2	0	1	0	0

V3	0	0	1	0
V4	0	0	1	0
V5	0	1	0	0
...
V156	1	0	0	0
V157	0	0	1	0
V158	0	1	0	0
V159	0	1	0	0
V160	0	0	1	0

F. Proses Clustering

Tabel 7. Menghitung Bobot Ed

No	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Total
V1	0,324	0,612	0,453	0,249	1,638
V2	0,509	1,130	0,385	0,214	2,239
V3	0,269	0,518	0,609	0,306	1,701
V4	0,275	0,543	0,622	0,304	1,743
V5	0,350	0,463	0,313	0,201	1,326
...
V156	0,634	0,609	0,289	0,182	1,713
V157	0,254	0,514	0,958	0,367	2,093
V158	0,333	0,606	0,427	0,241	1,606
V159	0,297	0,618	0,585	0,285	1,786
V160	0,239	0,446	0,787	0,372	1,843

1) Menghitung Bobot ED.

Perhitungan berdasarkan persamaan 4. Hasil pembobotan ED ditunjukkan pada tabel 7.

Perhitungan :

Diketahui :

Nilai ED = 3.085 \rightarrow atribut 1, responden 1

Maka : Bobot ED = $1 / 3.085 = 0.324 \rightarrow$ atribut 1, responden 1

2) Menghitung Derajat Keanggotaan

Perhitungan berdasarkan persamaan 5. Hasil derajat keanggotaan ditunjukkan pada tabel 8.

Perhitungan :

Diketahui :

Bobot ED = 0,324 \rightarrow atribut 1, responden 1

Total Bobot ED = 1,638

Maka : Derajat Keanggotaan = $0,324 / 1,638 = 0,198 \rightarrow$ atribut 1, responden 1

Tabel 8. Derajat Keanggotaan

No	Cluster 1	Cluster 2	Cluster 3	Cluster 4
V1	0,198	0,373	0,277	0,152
V2	0,228	0,505	0,172	0,096
V3	0,158	0,304	0,358	0,180
V4	0,158	0,311	0,357	0,174
V5	0,264	0,349	0,236	0,152
V156	0,370	0,355	0,169	0,106
V157	0,121	0,246	0,458	0,175
V158	0,207	0,377	0,266	0,150
V159	0,166	0,346	0,328	0,160
V160	0,130	0,242	0,427	0,202

3) Menghitung Kuadrat Derajat Keanggotaan

Perhitungan berdasarkan persamaan 6. Hasil kuadrat derajat keanggotaan ditunjukkan pada tabel 9.

Contoh Perhitungan :

Diketahui : Nilai derajat keanggotaan = 0,198 \rightarrow atribut 1, responden 1

Maka : Kuadrat derajat keanggotaan = $0,198^2 = 0,039 \rightarrow$ atribut 1, responden 1

Tabel 9. Kuadrat Derajat Keanggotaan

No	Cluster 1	Cluster 2	Cluster 3	Cluster 4
V1	0,039	0,139	0,077	0,023
V2	0,052	0,255	0,030	0,009
V3	0,025	0,093	0,128	0,032
V4	0,025	0,097	0,127	0,030
V5	0,069	0,122	0,056	0,023
...				
V156	0,137	0,126	0,029	0,011
V157	0,015	0,060	0,210	0,031
V158	0,043	0,142	0,071	0,022
V159	0,028	0,120	0,107	0,026
V160	0,017	0,058	0,182	0,041
Total	6,463	19,433	15,119	4,685

4) Menghitung bobot derajat keanggotaan

Perhitungan berdasarkan persamaan 7. Hasil bobot derajat keanggotaan ditunjukkan pada tabel 10.

Contoh Perhitungan :

Diketahui :

Kuadrat derajat keanggotaan = 0,039 \rightarrow atribut 1, responden 1

Feature vector = 4,000

Maka : Bobot derajat keanggotaan = $0,039 \times 4,000 = 0,157 \rightarrow$ atribut 1, responden 1

Tabel 10. Bobot Derajat Keanggotaan untuk keseluruhan Cluster

No	Cluster 1				
	C1	C2	C3	C4	C5
V1	0,157	0,130	0,104	0,078	0,143
V2	0,190	0,190	0,224	0,155	0,155
V3	0,050	0,058	0,075	0,075	0,092
V4	0,083	0,083	0,058	0,075	0,050
V5	0,208	0,278	0,162	0,324	0,232
...
V156	0,638	0,410	0,547	0,456	0,593
V157	0,044	0,034	0,034	0,044	0,034
V158	0,172	0,158	0,143	0,072	0,143
V159	0,074	0,101	0,092	0,083	0,055
V160	0,034	0,039	0,045	0,056	0,039
Total	20,784	20,683	20,055	21,515	20,687

No	Cluster 2				
	C1	C2	C3	C4	C5
V1	0,558	0,465	0,372	0,279	0,511
V2	0,934	0,934	1,104	0,765	0,765
V3	0,185	0,216	0,278	0,278	0,340
V4	0,323	0,323	0,226	0,291	0,194
V5	0,365	0,487	0,284	0,568	0,406
...
V156	0,589	0,379	0,505	0,421	0,547
V157	0,181	0,141	0,141	0,181	0,141
V158	0,568	0,521	0,474	0,237	0,474
V159	0,320	0,440	0,400	0,360	0,240
V160	0,117	0,136	0,156	0,195	0,136
Total	61,887	60,904	60,545	60,800	60,121

No	Cluster 3				
	C1	C2	C3	C4	C5
V1	0,306	0,255	0,204	0,153	0,281
V2	0,108	0,108	0,128	0,089	0,089
V3	0,256	0,299	0,384	0,384	0,469
V4	0,424	0,424	0,297	0,381	0,254
V5	0,167	0,223	0,130	0,260	0,186
...
V156	0,133	0,086	0,114	0,095	0,124
V157	0,629	0,489	0,489	0,629	0,489
V158	0,283	0,259	0,236	0,118	0,236
V159	0,286	0,394	0,358	0,322	0,215
V160	0,364	0,425	0,486	0,607	0,425
Total	43,947	42,735	44,099	43,812	43,029

No	Cluster 4				
	C1	C2	C3	C4	C5
V1	0,093	0,077	0,062	0,046	0,085
V2	0,034	0,034	0,040	0,027	0,027
V3	0,064	0,075	0,097	0,097	0,118
V4	0,101	0,101	0,071	0,091	0,061
V5	0,069	0,092	0,054	0,107	0,077
...
V156	0,053	0,034	0,045	0,038	0,049
V157	0,092	0,072	0,072	0,092	0,072
V158	0,090	0,082	0,075	0,037	0,075
V159	0,068	0,094	0,085	0,077	0,051
V160	0,081	0,095	0,108	0,136	0,095
Total	13,512	13,235	13,414	13,947	13,445

5) Pembangkitan Centroid Baru

Perhitungan berdasarkan persamaan 8. Hasil centroid baru ditunjukkan pada tabel 11.

Contoh perhitungan :

Diketahui :

Total bobot derajat keanggotaan = 20,784 → atribut 1
Total kuadrat derajat keanggotaan = 6,463 → atribut 1
Maka : Centroid baru = 20,784 / 6,463 = 3,216

Tabel 11. Centroid Baru

	<i>Attribut 1</i>	<i>Attribut 2</i>	<i>Attribut 3</i>	<i>Attribut 4</i>	<i>Attribut 5</i>
Cluster 1	3,216	3,200	3,103	3,329	3,201
Cluster 2	3,185	3,134	3,116	3,129	3,094
Cluster 3	2,907	2,827	2,917	2,898	2,846
Cluster 4	2,884	2,825	2,863	2,977	2,870

6) Perhitungan Nilai Error

Pada tahapan ini dimana dilakukan proses perhitungan nilai error yang diperoleh dari selisih antara centroid awal atau centroid sebelumnya dengan centroid baru yang telah dihitung. adapun hasil perhitungan ditunjukkan pada tabel 12.

Tabel 12. Perhitungan Nilai Error terhadap centroid awal dan centroid akhir

	<i>Atribut 1</i>	<i>Atribut 2</i>	<i>Atribut 3</i>	<i>Atribut 4</i>	<i>Atribut 5</i>	Total
Cluster 1	1,284	1,008	1,439	0,838	1,007	5,5759
Cluster 2	0,315	0,158	0,509	0,038	0,198	1,2185
Cluster 3	0,407	0,452	0,208	0,731	0,471	2,2691
Cluster 4	1,384	1,367	1,072	1,810	1,412	7,0447
Nilai Error						16,1082

Pada penelitian ini, proses *clustering* dihentikan pada iterasi ke 18 dikarenakan nilai *absolute* antara nilai centroid pada iterasi ke-17 dan nilai centroid pada iterasi ke-18 bernilai 0.000005 telah lebih kecil dari nilai error yang telah ditentukan yaitu 0,00001.

7) Pengelompokkan Data.

Selanjutnya data dapat dikelompokkan sesuai dengan jarak terdekat data pada iterasi ke-18. Hasil pengelompokkan data ditunjukkan pada tabel 13.

Tabel 13. Hasil Pengelompokkan

<i>CLUSTER</i>	<i>Nilai Centroid</i>	<i>Jumlah</i>	<i>Data Responden</i>
Program Studi D3 Teknik Informatika	3,0500	41	1,5,6,9,11,15,17,19,32,39,42,50,54,57,59,67,68,69,70,73,74,76,79,80,95,96,97,102,104,105,107,112,116,120,136,140,143,145,149,,151,156
	2,9979		
	3,0042		
	3,0688		
	2,9979		
Program Studi D3 Teknik Komputer	3,0500	46	2,4,7,14,18,22,24,25,28,30,31,34,36,40,41,44,46,48,60,62,63,64,65,66,77,78,84,85,91,92,94,103,109,111,115,119,126,127,128,135,138,141,144,147,158,159
	2,9979		
	3,0042		
	3,0687		
	2,9979		
Program Studi D4 Teknologi Informasi Multimedia	3,0500	21	20,21,23,26,33,43,56,71,83,99,101,106,117,121,124,125,132,133,150,153,157
	2,9979		
	3,0042		
	3,0688		
	2,9979		
Program Studi D4 Teknik Rekayasa Komputer	3,0500	52	3,8,10,12,13,16,27,29,35,37,38,45,47,49,51,52,53,55,58,61,72,75,81,82,86,87,88,89,90,93,98,100,108,110,113,114,118,122,123,129,130,131,134,137,139,142,146,148,152,154,155,160
	2,9979		
	3,0042		
	3,0688		
	2,9979		

G. Pengujian Hasil

Berdasarkan persamaan 9 diatas, hasil perhitungan sebagai berikut :

Contoh perhitungan :

Diketahui :

Nilai Atribut (Minat) Rsponden.1 = 4,000

Nilai Centroid K1 = 3,0500

$$\text{Maka : Absolute Values} = \frac{4,000-3,050}{4,000} = 0,237$$

Langkah selanjutnya menghitung persentase error MAPE. Hasil persentase error ditunjukkan pada tabel 14.

Contoh perhitungan :

Diketahui :

Nilai rata-rata semua atribut = 9,1613

Banyaknya data *cluster* 1 (n) = 41

Maka :

MAPE = 9,1613/ 41= 0,2234 jika dipersenkan hasil nya 22,34%

Tabel 14. Persentase Error

Mean Absolute Percentage Error (MAPE)	
CLUSTER	FUZZY C-MEANS
Program Studi D3 Teknik Informatika	22,34%
Program Studi D3 Teknik Komputer	22,75%
Program Studi D4 Teknologi Informasi Multimedia	34,55%
Program Studi D4 Teknologi Rekayasa Komputer	28,63%
	27,07%

4. KESIMPULAN

Klasifikasi minat siswa-siswa terhadap 4 Program Studi yaitu D3 Teknik Informatika, D3 Teknik Komputer, D4 Teknologi Informasi Multimedia dan D4 Rekayasa Teknologi Komputer yang dilakukan pada Jurusan Teknologi Informasi – Politeknik Negeri Samarinda. Hasil klasifikasi menggunakan metode Fuzzy C-means yaitu : Cluster 1 beranggotakan 41 siswa-siswi. Pada cluster 2 beranggotakan 46 siswa-siswi, Cluster 3 beranggotakan 21 siswa-siswi serta Cluster 4 beranggotakan 52 siswa-siswi. Dalam penelitian ini menggunakan metode fuzzy c-means, menggunakan batasan nilai errornya adalah 0.00001 dan perhentian perulangan iterasi terjadi dalam 18 kali iterasi. Rata-rata persentase MAPE untuk keseluruhan cluster sebesar 27.07%.

DAFTAR PUSTAKA

- [1] Karti, H.S., and Irhamah: 'Pengelompokan Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Indikator Pendidikan SMA/SMK/MA dengan Metode C-Means dan Fuzzy C-Means', *SAINS DAN SENI POMITS*, 2013, 2, (2), pp. 288 - 193
- [2] Wilandari, Y., Mukid2, M.A., M., N., and S., Y.A.: 'Analisis Klasifikasi Kabupaten Di Jawa Tengah Berdasarkan Populasi Ternak Menggunakan Fuzzy Cluster Means', *Media Statistika*, 2014, 7, (2)
- [3] Purnamasari, S.B., Yasin, H., and Wuryandari, T.: 'Pemilihan Cluster Optimum Pada Fuzzy C-Means (Studi Kasus: Pengelompokan Kabupaten/Kota Di Provinsi Jawa Tengah Berdasarkan Indikator Indeks Pembangunan Manusia)', *GAUSSIEN*, 2014, 3, (3), pp. 491 - 198
- [4] Ningrat, D.R., Maruddani, D.A.I., and Wuryandari, T.: 'Analisi Cluster Dengan Algoritma K-Means Dan Fuzzy C-Means Clustering Untuk Pengelompokan Data Obligasi Korporasi', *GAUSSIEN*, 2016, 5, (4), pp. 641-650
- [5] Wikarno, Malani, R., and Suprpty, B.: 'Perbandingan Metode K-Means dan Fuzzy C-Means Untuk Pengelompokan Pegawai Berdasarkan Nilai Kinerja dan Tingkat Kedisiplinan Pegawai ', *IEEE*, 2018
- [6] P.R. Nastiti, and A.B.W. Putra, "Perbandingan Algoritma K-Means dan Fuzzy C-Means Clustering Untuk Kualifikasi Data Kinerja Dosen di Jurusan Teknologi Informasi POLNES," *Prosiding Seminar SEBATIK*, Vol.1, No.1, pp.71-76, 2017
- [7] Islami, A.K., and Widodo, E.: 'Pengelompokan Kepemilikan Jaminan Kesehatan Menggunakan Metode Fuzzy C-Means Algorithm', *SI MaNIs*, 2017, 1, (1), pp. 299-305
- [8] Simbolon, C.L., Kusumastuti, N., and Irawan, B.: 'Clustering Lulusan Mahasiswa Matematika Fmipa UNTAN Pontianak Menggunakan Algoritma Fuzzy C-Means', *Bimaster*, 2013, 02, (1), pp. 21-26
- [9] Malani, and B. Suprpty, "Implementasi Sistem Seleksi Kinerja Pegawai Negeri Khususnya Tenaga Pengajar Terbaik Menggunakan Metode K-Means Clustering Pada Politeknik Negeri Samarinda," *Seminar Nasional Teknologi Terapan*, Vol.3, pp.1491, 2016